

周报

本周主要考虑了 palantir 的整体实现方案

1、global ontology 的构建

采用的方法有：

- 1) developer 手动构建，可利用 protégé
- 2) 利用半自动化抽取：从垂直或者百科网站上抽取相关信息（难度目前不好估计，如果找不到合适的网站进行抽取将不可行）
- 3) user 在界面上手动构建，也可修改现有的 ontology

2、local ontology 的构建

1) 用户导入结构化的 data table。然后用户手动选择每一列的相关内容，
developer 根据 user 的标注构建 local ontology

- 2) 对于非结构化的文件，完全由用户手动构建

3、整体知识图的构建

- 1) 由用户导入的结构化的 data table
- 2) 用户对于非结构化文件手动构建的知识
- 3) 半自动化知识的构建。采用从垂直网站上抽取知识的方法（可行性较大，应该不会太难，之前已经上已经实现过）

4、object model 的构建

程序中的 data model 目前考虑的有以下几种：

- 1) DataSources
- 2) Objects（可以再细分为 Document, entities, events）
- 3) Properties
- 4) Notes
- 5) Media
- 6) Links（也就是 relationship）
- 7) Data Source Record